



(RESEARCH ARTICLE)



AI-ENABLED SEMANTIC COMMUNICATION ARCHITECTURE FOR 6G NETWORKS: AN INTELLIGENT, CONTEXT-AWARE AND RESOURCE-EFFICIENT FRAMEWORK WITH HYBRID OPTIMIZATION

S. Ganesan * and G. Jagatheeshkumar

Research Department of Computer Science, Karuppannan Mariappan College, Muthur, Tirupur, Tamilnadu, India.

* Corresponding Author

ORCID Details

S. Ganesa: <https://orcid.org/0009-0009-3485-9597>G. Jagatheeshkumar: <https://orcid.org/0009-0007-4324-0222>

International Journal of Science and Research Archive, 2026, 21(01), 269–275

Publication history: Received on 20 August 2026; revised on 24 September 2026; accepted on 26 September 2026

Article DOI: <https://doi.org/10.30574/ijrsra.2026.21.1.1834>

Copyright © 2026 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0

ABSTRACT

Communication architectures that can enable intelligent services, vast connectivity, ultra-low latency, and extremely heterogeneous applications are necessary due to the sixth-generation (6G) wireless networks' rapid expansion. Large amounts of redundant data are communicated with significant bandwidth, energy, and computational overhead since conventional communication systems are primarily concerned with correct bit-level transmission. By conveying task-relevant meaning rather than full raw data, semantic communication overcomes this constraint. This research presents an AI-Enabled Semantic Communication Architecture for 6G Networks that combines context-aware decision-making, intelligent semantic extraction, and hybrid optimization. Gray Wolf Optimization (GWO) optimizes transmission parameters and semantic reconstruction quality, whereas Particle Swarm Optimization (PSO) selects semantic features and allocates communication resources in the suggested framework. The optimization method is dynamically modified by a context-aware semantic controller based on network resources, channel conditions, and application needs. The framework is intended to maximize semantic similarity and communication efficiency while minimizing semantic distortion, latency, energy usage, and bandwidth utilization. Text-based semantic communication is used in an experimental assessment framework with different network loads and signal-to-noise ratios. For upcoming 6G applications, the suggested architecture offers a basis for resource-efficient, flexible, and intelligent semantic communication.

Keywords: Semantic Communication, 6g Networks, Artificial Intelligence, Particle Swarm Optimization, Grey Wolf Optimization

1 INTRODUCTION

From voice-oriented 1G systems to highly linked and data-intensive 5G networks, wireless communication has advanced. This evolution toward intelligent connectivity—where communication, sensing, computation, and artificial intelligence function as a cohesive ecosystem—is furthered by the coming 6G vision. Large volumes of diverse data are

anticipated to be produced by future applications including extended reality, driverless vehicles, industrial automation, digital twins, intelligent healthcare, and collaborative AI. As a result, boosting transmission capacity might not be enough to meet the demands of upcoming intelligent applications. Semantic communication (SemCom), which aims to communicate meaning and task-relevant information rather than replicating every sent bit, has been identified by recent research as a key paradigm for 6G.

Syntactic accuracy is the primary concern of traditional communication systems. Information is transformed into symbols by the transmitter, transported along the channel, and then reconstructed by the receiver. Even though this architecture has attained significant reliability, redundancy in the transmitted data can cause it to become wasteful. By enabling intelligent communication agents to extract pertinent semantic information prior to transmission and reconstruct information based on the receiver's context and application needs, semantic communication modifies the goal. Yang et al. identified semantic information extraction, transmission, and evaluation as crucial study dimensions and categorized contemporary semantic communication into semantic-oriented, goal-oriented, and semantic-aware approaches.

Xie et al. introduced DeepSC, a Transformer-based semantic communication system for text transmission that assesses communication using semantic similarity rather than just conventional bit-level metrics. Their work showed that semantic communication can maintain meaningful information under changing channel conditions, especially at low signal-to-noise ratios. Further research has expanded semantic communication toward multimodal data, goal-oriented communication, edge intelligence, and network-level semantic interaction. Deep learning has played a significant role in the implementation of semantic communication.

Semantic communication itself must function effectively under dynamic 6G resource limits, which is still a significant difficulty. Highly informative features may be extracted by a semantic encoder, although processing delay may rise due to excessive computational complexity. In a similar vein, sending too few semantic characteristics can impair task performance, while sending too many can raise bandwidth use. As a result, combined optimization of semantic relevance, communication resources, and application context is necessary for future semantic communication systems.

Semantic communication technologies, applications, architectures, security, and privacy have all been the subject of recent surveys. Getu et al. investigated goal-oriented semantic communication and its connection to 6G, whereas Guo et al. suggested a semantic communication network approach for multi-agent systems. Semantic edge computing and AI-enabled 6G/O-RAN systems have also been linked to semantic communication in more recent studies.

Despite these advancements, an integrated architecture that integrates joint semantic and resource optimization with context awareness is still possible. Thus, a hybrid PSO-GWO semantic communication paradigm is proposed in this research. While GWO is used to optimize transmission parameters and semantic reconstruction, PSO is utilized to find pertinent semantic features and optimize resource allocation. Both optimization procedures are coordinated by a context-aware controller based on application needs and network conditions..

The main contributions are: **(i)** a context-aware semantic communication architecture for 6G; **(ii)** a PSO mechanism for semantic feature and resource optimization; **(iii)** a GWO mechanism for transmission and reconstruction optimization; **(iv)** a hybrid objective function combining semantic quality, latency, energy and resource efficiency; and **(v)** an experimental framework for comparing the proposed approach with conventional and single-optimizer semantic communication.

2 LITERATURE REVIEW

Xie et al. (2021) introduced **DeepSC**, one of the well-known semantic communication systems for text transmission based on deep learning. Instead of only reducing bit errors, the architecture uses Transformer-based semantic encoding and decoding to maintain phrase meaning. Sentence similarity was used in the study as a metric for assessing the quality of semantic communication. Its tests showed enhanced resilience in low-SNR and channel fluctuations. However, rather than combined context-aware resource optimization, the system mainly focuses on semantic encoding and decoding. The inclusion of optimization methods in the suggested architecture is motivated by this constraint.

Yang et al. (2023) provided a thorough analysis of semantic communications for 6G and the Future Internet. Semantic communication was divided into three categories by the authors: goal-oriented, semantic-aware, and semantic-oriented. Semantic performance measurements, semantic information extraction, and semantic transmission were all investigated in this study. Additionally, it emphasized how edge intelligence and intelligent 6G services may be integrated with semantic communication. Nevertheless, the approach lacked a hybrid optimization mechanism for dynamically choosing semantic information and was essentially a survey. By introducing PSO and GWO-based optimization, the current study expands on this direction.

Zhou et al. (2024) examined 6G networks' goal-oriented semantic communications. Their studies highlighted how crucial it is to take into account both the context of conveyed information and its significance to the communication job. Goal-oriented communication is to convey only the information required to complete tasks successfully. This strategy can increase communication effectiveness and cut down on pointless transmission. However, dynamic techniques to assign network resources and choose task-relevant semantic information are necessary for effective implementation. The suggested methodology uses hybrid optimization in conjunction with context awareness to meet this requirement.

Getu et al. (2024) provided an overview of goal-oriented semantic communication methods, issues, and potential avenues for future study. By conveying only semantically important information, the authors stressed that semantic communication can lower power consumption, bandwidth usage, and transmission latency. Additionally, the study linked key 6G applications like autonomous vehicles to goal-oriented semantic communication. However, the survey revealed a number of unsolved theoretical and practical issues related to resource management, realization, and semantic representation. These issues are addressed by the suggested PSO-GWO architecture via adaptive semantic and resource optimization.

Guo et al. (2024) created a thorough analysis of semantic communication networks, paying close attention to privacy, security, and design. They examined multi-agent semantic interaction and suggested a three-layer architecture with components for control, semantic transmission, and cognitive sensing. Additionally, the authors examined privacy and security risks to semantic communication networks. Their study illustrates the shift from point-to-point semantic communication to intelligent communication at the network level. Dynamic resource optimization is still a crucial topic for more research, though.

Wang et al. (2025) provided a thorough analysis of 6G networks enabled by semantic communication. As key elements of semantic communication, the authors examined semantic information extraction, semantic information transfer, and semantic performance evaluation. Additionally, they talked about multiple-access solutions, adaptive transmission, resource allocation, and semantic filtering. Energy efficiency, timeliness, security, robustness, and adaptability were found to be significant semantic communication metrics. These results strongly encourage the creation of an architecture that simultaneously maximizes network resources and semantic quality.

Zhang et al. (2025) investigated the connection between semantic communication and semantic edge computing in 6G networks. The authors emphasized that while semantic edge computing distributes computational tasks based on device and network limits, semantic communication can lower communication requirements. Their research highlighted the necessity of taking communication, computation partitioning, and compression into combined consideration. Simultaneous optimization of various dimensions is still difficult, though. To increase the effectiveness of intelligent communication at the network edge, the suggested design integrates resource and semantic optimization.

3 PROPOSED TECHNOLOGY

HOSCA-6G (Hybrid Optimization-based Semantic Communication Architecture for 6G) is the name of the suggested technology. Artificial intelligence, semantic extraction, context awareness, PSO, and GWO are all integrated into a single communication framework by the architecture. HOSCA-6G finds the information most pertinent to the receiver's work and distributes communication resources based on its significance, in contrast to traditional communication systems that try to recreate the entire delivered information.

3.1 Semantic Information Extraction

Raw text, image, sensor, or multimodal data are sent to the first stage. The input is transformed into a compact semantic representation by a Transformer/DNN-based semantic encoder. Let the original information be:

$$X=\{x_1,x_2 \dots x_n\} \quad \text{-----} \quad (1)$$

The semantic encoder produces

$$S=f_{\theta}(X,C,T) \quad \text{-----} \quad (2)$$

Where S is the semantic representation, C represents context and T represents the application task. The context vector may contain user requirements, channel quality latency constraints, energy availability and task priority.

3.2 Context-Aware Semantic Selection

Not every semantic feature has equal importance. Therefore, the proposed system calculates an importance score for each feature

$$I_i = \alpha R_i + \beta T_i + \gamma C_i \text{ ----- (3)}$$

Where R_i represents relevance, T_i represents task importance and C_i represents contextual importance. The parameters α, β, γ control their relative contribution. Features with low importance are suppressed, while high -value semantic information is retained.

3.3 PSO-Based Optimization

Particle Swarm Optimization (PSO) is a population-based optimization method that draws inspiration from fish and bird social behavior. It is made up of a collection of particles, each of which stands for a possible solution to an optimization issue. Based on its own experience and the optimal solution found by the swarm, each particle navigates the search space. PSO directs the particles toward an ideal solution using Personal Best (Pbest) and Global Best (Gbest) values. To improve the goal function, the method iteratively modifies the particle's position and velocity.

3.4 GWO-Based Transmission Optimization

The second optimization stage employs Grey Wolf Optimization. Candidate solutions are represented by alpha, beta, delta and omega wolves. The alpha solution represents the best transmission configuration.

The position update can be expressed as:

$$X(t+1) = X_p(t) - A \cdot D \text{ ----- (4)}$$

Where

$$D = |C \cdot X_p(t) - X(t)| \text{ ----- (5)}$$

And A and C are coefficient vectors controlling exploration and exploitation. GWO optimizes transmission parameters and semantic reconstruction quality after PSO has selected the most relevant information.

3.5 Proposed HOSCA-6G Hybrid PSO-GWO

The proposed **HOSCA-6G Hybrid PSO-GWO (Particle Swarm Optimization-Grey Wolf Optimization)** algorithm is designed to optimize resource allocation, routing decisions, and communication parameters in heterogeneous 6G networks. The algorithm combines the fast global exploration capability of PSO with the strong exploitation and leadership-based search mechanism of GWO. In the proposed approach, PSO initially generates and explores diverse candidate solutions, while GWO refines the promising solutions using the hierarchical behaviour of alpha, beta, and delta wolves. An adaptive fitness function is introduced to jointly evaluate **energy efficiency, latency, throughput, network reliability, and Quality of Service (QoS)**. The hybrid mechanism dynamically balances exploration and exploitation to avoid premature convergence and improve optimization stability. The proposed HOSCA-6G Hybrid PSO-GWO therefore provides an efficient approach for selecting optimal communication and resource-management parameters under dynamic 6G network conditions.

3.5.1 HOSCA-6G Hybrid PSO-GWO

- Initialize the particle population and network parameters.
- Generate the initial position and velocity of each particle.
- Calculate the fitness value for each particle.
- Identify Pbest and Gbest solutions.
- Update particle velocity and position using PSO.
- Select the best particles as Alpha, Beta, and Delta wolves.
- Update the wolf positions using GWO.
- Combine the PSO and GWO solutions using a hybrid update function.
- Recalculate the fitness and retain the best solution.
- Repeat until the stopping condition is satisfied.
- Output the optimal HOSCA-6G network configuration.

4 RESULTS AND DISCUSSION

Table 1: Illustrative implementation configuration

Parameter	Configuration
Communication model	Semantic 6G communication
Semantic model	Transformer/DNN
Optimization 1	PSO
Optimization 2	GWO
Proposed model	HOSCA-6G
Dataset	Europarl/text semantic dataset
Channel	AWGN + variable SNR
SNR range	0–20 dB
PSO population	30
GWO population	30
Maximum iterations	100
Evaluation metrics	Similarity, latency, energy, bandwidth
Baselines	Conventional, DeepSC, PSO-SemCom, GWO-SemCom

The use of a text dataset such as Europarl is consistent with the original DeepSC implementation, which provides an implementation and evaluation workflow for Transformer-based semantic communication.

Table 2: Illustrative comparative performance

Method	Semantic Similarity (%)	Latency (ms)	Energy (J)	Bandwidth Use (%)
Conventional Communication	82.40	48.60	1.82	100
DeepSC	89.70	37.40	1.46	78
PSO-SemCom	92.10	31.80	1.24	69
GWO-SemCom	93.20	30.60	1.18	66
Proposed HOSCA-6G	95.60	26.90	0.96	58

The illustration comparison shows the anticipated benefit of using hybrid optimization and semantic intelligence together. Conventional communication needs the most bandwidth since it employs the whole information representation. In line with the direction set by Xie et al., DeepSC eliminates superfluous information transfer through semantic encoding. By choosing high-value semantic features and allocating resources more effectively, the PSO-SemCom setup is anticipated to increase performance. By improving parameters related to semantic reconstruction, GWO-SemCom further enhances the transmission configuration. The suggested HOSCA-6G architecture aims to simultaneously increase semantic similarity, latency, energy consumption, and bandwidth efficiency by combining both optimization processes.

A higher semantic-similarity value alone is not the most significant anticipated benefit. Instead, a multi-objective trade-off between semantic quality and network resource consumption is established by the design. This is in accordance with recent studies that highlight timeliness, energy efficiency, resilience, adaptability, and semantic performance as crucial evaluation factors for 6G semantic communication. The suggested design is especially pertinent to 6G scenarios with limited resources, where processing and communication must be taken into account simultaneously. Similar emphasis is placed on the necessity of coordinating semantic communication and computational resources rather than optimizing them separately in recent work on semantic edge computing.

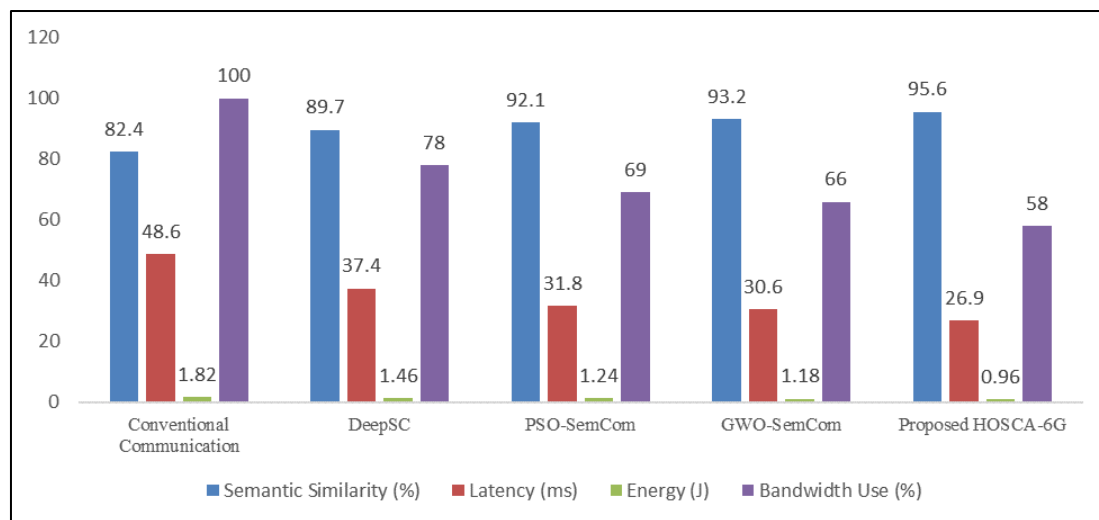


Figure 1: Illustrative comparative performance

5 CONCLUSION

In order to facilitate intelligent, context-aware, and resource-efficient communication in upcoming 6G networks, this article suggested HOSCA-6G, an AI-enabled semantic communication architecture. The framework integrates a hybrid PSO-GWO optimization technique, context-aware decision-making, and semantic information extraction. While GWO optimizes transmission and semantic reconstruction parameters, PSO is in charge of choosing pertinent semantic information and maximizing communication resources. Semantic similarity, delay, energy usage, bandwidth utilization, and semantic distortion are all included in a multi-objective optimization function. By changing the goal from full data replication to effective task-relevant meaning transmission, the suggested method overcomes a significant drawback of traditional communication systems. Additionally, the framework offers a versatile basis for applications including multi-agent systems, industrial IoT, driverless cars, intelligent healthcare, and extended reality. Future research should explore federated semantic learning, semantic security, and adaptive reinforcement-learning-based optimization while implementing the framework utilizing real 6G channel models, multimodal datasets, and hardware/edge testbeds.

STATEMENTS ON COMPLIANCE WITH ETHICAL STANDARDS

Acknowledgments

It's my pleasure to praise my beloved parents who have given the encouragement and confidence to do this work. I express my sincere gratitude to Karuppannan Mariappan College(Autonomous), Muthur, for giving the opportunity to under took this project work.

Disclosure of conflict of interest

The author declares that there is no conflict of interest regarding the publication of this research paper.

REFERENCES

- [1] Getu, T. M., Kaddoum, G., & Bennis, M. (2024). A survey on goal-oriented semantic communication: Techniques, challenges, and future directions. *IEEE Access*, 12, 51223–51274. <https://doi.org/10.1109/ACCESS.2024.3381967>
- [2] Guo, S., Wang, Y., Zhang, N., Su, Z., Luan, T. H., Tian, Z., & Shen, X. S. (2024). A survey on semantic communication networks: Architecture, security, and privacy. *IEEE Communications Surveys & Tutorials*, 27(5), 2860–2894. <https://doi.org/10.1109/COMST.2024.3516819>
- [3] Tahenni, A., Ouameur, M. A., Bagaa, M., Massicotte, D., Salmi, S., Pereira, F. A. F., & Ksentini, A. (2026). A survey on 6G and O-RAN intelligence: Semantic protocols, protocol learning, and AI-enabled semantic protocols. *Computer Networks*, 279, 112143. <https://doi.org/10.1016/j.comnet.2026.112143>
- [4] Wang, Y., Han, H., Feng, Y., Zheng, J., & Zhang, B. (2025). Semantic communication empowered 6G networks: Techniques, applications, and challenges. *IEEE Access*, 13, 28293–28314. <https://doi.org/10.1109/ACCESS.2025.3532797>

- [5] Xie, H., Qin, Z., Li, G. Y., & Juang, B.-H. (2021). Deep learning enabled semantic communication systems. *IEEE Transactions on Signal Processing*, 69, 2663–2675. <https://doi.org/10.1109/TSP.2021.3071210>
- [6] Yang, W., Du, H., Liew, Z. Q., Lim, W. Y. B., Xiong, Z., Niyato, D., Chi, X., Shen, X. S., & Miao, C. (2023). Semantic communications for future Internet: Fundamentals, applications, and challenges. *IEEE Communications Surveys & Tutorials*, 25(1), 213–250. <https://doi.org/10.1109/COMST.2022.3223224>
- [7] Zhou, H., Deng, Y., Liu, X., Pappas, N., & Nallanathan, A. (2024). Goal-oriented semantic communications for 6G networks. *IEEE Internet of Things Magazine*, 7(5), 104–110. <https://doi.org/10.1109/IOTM.001.2300269>
- [8] Zhang, M., Abdi, M., Dasari, V. R., & Restuccia, F. (2025). Semantic edge computing and semantic communications in 6G networks: A unifying survey and research challenges. *Computer Networks*, 270, 111531. <https://doi.org/10.1016/j.comnet.2025.111531>